

Distributed Spectrum Sensing with Sequential Ordered Transmissions to a Cognitive Fusion Center

Laila Hesham, Ahmed Sultan and Mohammed Nafie

Wireless Intelligent Networks Center (WINC)

Nile University, Cairo, Egypt.

E-mail: laila.hesham@nileu.edu.eg, {asultan, mnafie}@nileuniversity.edu.eg

Abstract—Cooperative spectrum sensing is a robust strategy that enhances the detection probability of primary licensed users. However, a large number of detectors reporting to a fusion center for a final decision causes significant delay and also presumes the availability of unreasonable communication resources at the disposal of a network searching for spectral opportunities. In this work, we employ the idea of sequential detection to obtain a quick, yet reliable, decision regarding primary activity. Local detectors take measurements, and only a few of them transmit the log likelihood ratios (LLR) to a fusion center in descending order of LLR magnitude. The fusion center runs a sequential test with a maximum imposed on the number of sensors that can report their LLR measurements. We calculate the detection thresholds using two methods. The first achieves the same probability of error as the optimal block detector. In the second, an objective function is constructed and decision thresholds are obtained via backward induction to optimize this function. The objective function is related directly to the primary and secondary throughputs with inbuilt privilege for primary operation. Simulation results demonstrate the enhanced performance of the approaches proposed in this paper. We also investigate the case of fading channels between the local sensors and the fusion center, and the situation in which the sensing cost is negligible.*

Index Terms—Cognitive Radios, Cooperative Spectrum Sensing, Detection delay, Sequential detection.

I. INTRODUCTION

Cognitive radio is an emerging technology aimed at solving the problem of spectrum under-utilization caused by static spectrum allocation [1]. The technology allows a cognitive, also called unlicensed or secondary, terminal to make efficient use of any available spectrum at any given time. This involves the detection of the activity of licensed, or primary, users so that the secondary operation does not interfere with the primary networks and disrupt their services. This detection task may be extremely difficult due to a large array of factors including, inter alia, the wide variety of primary networks and the uncertainty about the propagation conditions between an active primary terminal and a secondary sensor attempting to detect primary activity. If the channel between a primary transmitter and a secondary sensor is in deep fade, the sensor would falsely decide that the probed channel is vacant. Secondary transmission on this channel would then cause interference on

the primary link. Relying on one sensor to detect the presence of primary operation is therefore highly unreliable [2].

Enhancement of sensing reliability requires a system of spatially distributed multiple sensors cooperating together in order to mitigate the impact of channel uncertainty provided that the channels between them and the primary transmitter are independent. Many papers have shown improvement in spectrum sensing capabilities through cooperation between individual cognitive users (see, e.g., [3], [4], [5], [6], [7]). However, in order to reap the benefits of distributed detection and cooperative diversity, efficient schemes are needed to combine the data obtained by the local sensors [8]. There is typically a fusion center at which the final decision is made. Whether the fusion center receives the raw unprocessed data from local sensors or summary messages obtained after some local processing, there are challenges concerning the transfer of data to the fusion center. It may be assumed, for instance, that there are orthogonal channels available for the transmission from the local detectors to the fusion center. In the context of cognitive radio, this assumption might not be practical given that the main objective of the cognitive users is to find communication resources. Moreover, if the duration of data transmission between the local sensors and the fusion center is large, then the primary user may switch back to activity, thereby depriving the secondary network of a communication opportunity. In the cognitive radio setting, not only many terminals are needed for robust detection of primary activity, but also this detection should be done in the quickest way possible and under very strict constraints regarding the communication between the cognitive detectors and the fusion center.

Sequential detection is known to reduce the time of detection for a certain specified detection reliability (see, e.g., [9], [10], [11]). In contrast with block detection that operates on a fixed predetermined number of samples, the number of samples in sequential detection is a random variable. In addition, for the case of binary hypothesis testing, sequential detection requires two thresholds and not only one for operation. This allows for the decision to continue acquiring more samples to satisfy some detection performance measures. There has been considerable interest in applying sequential analysis to distributed or decentralized detection problems (see, e.g., [12], [13], [14], [15], [16], [17]). According to [18], distributed implementation of sequential detectors can

*Part of this work was presented in ICASSP, May 2011.

This work was supported in part by a grant from the Egyptian NTRA (National Telecommunications Regulatory Authority).

have several forms. For instance, the sensors may forward their local decisions to the fusion center, which then runs a sequential test to obtain a global decision. On the other hand, the sensors themselves may perform sequential tests and cooperate together with or without a fusion center. In [19], the problem of distributed sequential detection in the presence of communication constraints is studied. An algorithm is developed for optimal rate or bandwidth distribution among detectors under a fixed bandwidth constraint. The work is based on the sequential probability ratio test (SPRT) of Wald [9].

In the context of cognitive radio, distributed sequential detection has been investigated, for example, in [20], [21], [22], [23]. In the scheme proposed in [21], each sensor computes the log-likelihood ratio (LLR) for its observations and reports it to the fusion center over a perfect reporting channel. The LLR's are accumulated sequentially at the fusion center till their sum is found sufficient to cross either of two predefined thresholds. The thresholds are obtained on the basis of the desired global detection and false alarm probabilities. The authors also investigate the case of sequential detection with model uncertainties. In [22], the detectors are divided into different sets according to the local signal-to-noise (SNR) of the sensing channel. The set with highest SNR sends to the fusion center first, followed by lower SNR until the fusion center can make a final decision. A doubly sequential test is proposed in [23] in which both the local detectors and the fusion center operate sequential tests. The detectors have thresholds computed using Wald's approximations [9]. When the first sensor reaches a decision, it conveys its decision to the fusion center which is running a second sequential test for a global decision on channel availability. The authors make several suggestions pertaining to the termination criterion at the fusion center. For example, the fusion center stops when it has received a specified number of decisions in favor of a particular hypothesis.

In addition to distributed sequential detection, we make use of the idea of ordered transmissions provided in [24] for the purpose of energy efficient signal detection in wireless sensor networks (see also [25]). By ordered transmissions, it is meant that the local sensor with the most reliable current measurement sends first its LLR value to the fusion center. In order for a sensor to know that it has the most reliable LLR value, there is no need for extensive information exchange. A timer backoff mechanism is suggested in [24] to resolve this issue. Note that the sensor selection criterion is its instantaneous measurement quality and not the average SNR as in [22] for example. We discuss the scheme in detail in this paper. Suffice it now to mention that the proposed scheme in [24] achieves the minimum average probability of error attained when all the sensor LLR's are used for a decision, and at the same time lowers the average number of transmissions needed to reach a decision. The scheme in [24] is utilized in [26] for sequential fusion in wireless sensor networks. It is certainly beneficial for a cognitive radio setting because it enhances the secondary throughput by quickly seizing the available spectral

opportunities.

In this paper, in order to reach a global decision at the secondary fusion center as quickly and reliably as possible, we adopt the ordered transmission scheme of [24]. However, we impose a strict constraint on the maximum number of transmissions to the fusion center. We therefore extend the scheme to work for the case when the fusion center is to make a decision after receiving a maximum of K observations from a total of M local sensors. Assuming a slotted primary network where the primary user may switch activity every fixed amount of time, only few observations can be used within a time slot to make a decision regarding the availability of a transmission opportunity. In other words, the decision regarding primary activity must be made soon enough to allow for channel access, and under the reasonable assumption of a narrowband cognitive control channel, only a few sensors can actually participate in the detection process.

Having imposed this constraint of a maximum number of reports from the local sensors, we devise a scheme where the decision thresholds employed at the secondary fusion center are computed using a dynamic programming framework. Given the LLR observations sent to the fusion center, the posterior probability of the channel being free is calculated and compared to the thresholds for a decision. We derive expressions for the likelihood functions of ordered LLR's and use these in both proposed schemes. For the dynamic programming scheme, a weighted sum of primary and secondary throughput objective function is constructed that also accounts for transmission costs, lost transmission opportunities and a penalty for collision with primary transmission. It is important to note that, in [27], dynamic programming is also employed to get the decision thresholds. However, it is assumed that all cognitive sensors would be able to report their energy measurements to the fusion center, where each detector senses one of the multi-band channels.

In summary we make the following contributions in this paper. We use a cooperative sequential detection framework to address the tradeoff between throughput and detection reliability in cognitive radio systems taking into account the strict time limitation during which the decision must be made, which also imposes a limit on the number of sensors involved in the decision process. We employ an ordered LLR transmission scheme modifying the method of threshold determination in [24] to account for the constraint on the number of reporting sensors. We also use dynamic programming to obtain the decision thresholds. The proposed method improves the throughput performance, which is the desired goal in the context of cognitive radios. We prove that if sensing has no cost, the optimal sequential decision is either to declare that the primary is idle or to continue sampling without ever deciding that the primary is active except possibly at the last decision stage. Finally, we address the impact of fading on the reporting channels linking the cognitive detectors to the fusion center.

The rest of the paper is organized as follows: Section II presents the system model of cooperative spectrum sensing in a cognitive radio network. Section III discusses the previous

work on ordered transmissions scheme, then discusses our proposed extension derivation. Section IV and V describe the proposed dynamic programming scheme. The effect of fading on the channels between the local detectors and the fusion center is explained in Section VI. In Section VII we provide simulation results and elaborate on the difference in performance when both schemes are used. We conclude the paper in Section VIII.

II. SYSTEM MODEL

We consider a slotted primary system as shown in Figure 1, where the primary activity, whether on or off, does not change during the time slot duration, τ_s . Primary activity switches independently from one slot to the next. There are M cognitive sensors which take a number of measurements at the beginning of each time slot and compute a function of these measurements. A maximum of K sensors among the M sensors forward the results sequentially to a fusion center at which the final decision regarding primary activity is taken. We consider binary hypothesis testing at the fusion center with the following two hypotheses:

H_0 : Sensed channel is free ; H_1 : Sensed channel is busy

The prior probabilities of each, denoted by π_0 and $(1 - \pi_0)$, respectively, are assumed to be known. Observations from different sensors are conditionally independent given either hypothesis but can be non-identical. Let $X_i(n)$ be the received signal at the i^{th} sensor at instant n , where $i = 1, 2, \dots, M$. At each sensor i , $X_i(n)$ are independent given each hypothesis and are identically distributed. Assume that a total of N samples are taken over a time duration τ_N . Under the two hypotheses, $X_i(n)$ is given by

$$\begin{aligned} H_0 : X_i(n) &= W_i(n), & n &= 1, 2, \dots, N \\ H_1 : X_i(n) &= S_i(n) + W_i(n), & n &= 1, 2, \dots, N \end{aligned} \quad (1)$$

where W_i is additive white Gaussian noise (AWGN) having the same noise power, σ^2 , at all sensors. Without loss of generality, the received primary signal $S_i(n)$ is assumed to be real zero-mean Gaussian random variable. The conditional probability distributions of $X_i(n)$ given H_0 and H_1 are described by $f_{X_i(n)}(x_n | H_0)$ and $f_{X_i(n)}(x_n | H_1)$, respectively, such that

$$\begin{aligned} f_{X_i(n)}(x_n | H_0) &\sim N(0, \sigma^2) \\ f_{X_i(n)}(x_n | H_1) &\sim N(0, \sigma_{s_i}^2 + \sigma^2) \end{aligned} \quad (2)$$

where $\sigma_{s_i}^2$ is defined as the average received primary signal power at the i^{th} local sensor, and is assumed to be fixed over a time slot and to change relatively slowly over time. The values of $\sigma_{s_i}^2$ are assumed to be known at the local sensors and are transmitted periodically after being quantized and coded to the fusion center on a low-rate common control channel that is not on the band being sensed. The quantization noise effect is assumed to be negligible. The LLR at the i^{th} sensor is defined as

$$Y_i = \sum_{n=1}^N \log \left[\frac{f_{X_i(n)}(x_n | H_1)}{f_{X_i(n)}(x_n | H_0)} \right] \quad (3)$$

This is the quantity that is computed by the local detector and reported to the fusion center sequentially as explained below. Defining the local signal-to-noise ratio (SNR) as $\gamma_i = \frac{\sigma_{s_i}^2}{\sigma^2}$, we can compute the LLR at the i^{th} sensor as

$$\begin{aligned} Y_i &= \frac{1}{2\sigma^2} \cdot \frac{\gamma_i}{\gamma_i + 1} \sum_{n=1}^N |X_i(n)|^2 \\ &\quad - \log(1 + \gamma_i) \frac{N}{2} \end{aligned} \quad (4)$$

Note that under H_0 the summation $\sum_{n=1}^N \frac{|X_i(n)|^2}{\sigma^2}$ is a chi-square distribution with N degrees of freedom. This is the same for the summation $\sum_{n=1}^N \frac{|X_i(n)|^2}{\sigma^2(1+\gamma_i)}$ under H_1 . That is, the likelihood functions of Y_i given H_0 or H_1 are shifted and scaled chi-square distributions with N degrees of freedom.

The LLR values are quantized and coded using a fixed number of bits. As in the case of the local signal power, the quantization noise is assumed to be negligible. The information exchange between the cognitive sensors and the fusion center occurs on the common control channel. We initially assume that the transmission of the LLR values to the fusion center is perfect, then we consider the effect of fading in the channels between the local sensors and the fusion center in Section VI. We assume that the procedure of seizing the control channel and sending the observation of one sensor requires an amount of time which we denote as τ . Figure 1 illustrates the time durations τ_N , τ and τ_s .

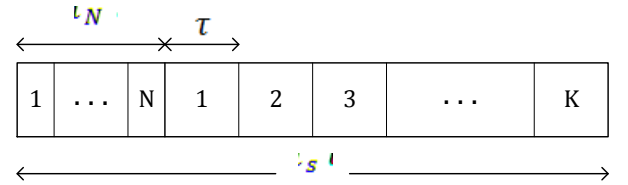


Fig. 1. The primary time slot has a duration of τ_s units of time. Over τ_N , each one of M cognitive sensors takes N samples at the beginning of the time slot. Afterwards, there are K mini-slots, each of duration τ , over which one LLR observation is transmitted from one cognitive sensor to the fusion center over a low-rate common control channel.

Since the transmission of each computed LLR value takes some time τ , eliciting another measurement from the sensors, though improving the reliability of detection, wastes a duration of τ from the potential secondary transmission opportunity in case the primary is off. This causes a decrease in secondary throughput. In other words, we have a reliability-throughput tradeoff [28] which we can control by allowing only the subset of the cognitive detectors with the most reliable observations to transmit their LLR's to the fusion center. This is implemented

by having two thresholds at the fusion center. The LLR with maximum magnitude is transmitted first to the fusion center. A decision can be made, but if the metric is between the two thresholds, the second highest LLR in magnitude is transmitted to the fusion center and is combined with the first, and then a decision is attempted again. This continues until a decision in favor of H_0 or H_1 is made, or K LLR's are accumulated at the fusion center. If k sensors, $1 \leq k \leq K$, are probed before a decision is reached, the time to make a decision is $\tau_N + k\tau$. Since the slot duration is τ_s , this leaves $\tau_s - \tau_N - k\tau$ for transmission. Given the durations τ_N , τ and τ_s , it is obvious that K must satisfy the inequality $\tau_s - \tau_N - K\tau \geq 0$.

Some final notes are in order regarding the operation of the secondary network.

- (a) The LLR values can be quantized to a few bits only with no considerable performance degradation when compared with the case where unquantized LLR's are used [29].
- (b) We assume a single hop network in which any broadcast by a secondary node is heard by the fusion center and all other nodes. This assumption is necessary such that when a node seizes the common channel to transmit its LLR, all other nodes become aware of its transmission.
- (c) We assume that the sensors have nearly synchronized clocks. This issue and its practical aspects are discussed extensively in [30].

III. ORDERED TRANSMISSIONS SCHEME

In this section we describe the scheme in [24]. A backoff timer is set at each of the M sensors according to the magnitude of the locally computed LLR. Specifically, the timer is decreased as the absolute value of the sensor's LLR increases. Consequently, transmissions proceed such that the most confident observations and informative measurements are sent first to the fusion center. The fusion center then compares the accumulated sum of the received LLR's to two thresholds and makes one of three decisions accordingly; declare H_0 , continue taking the next ranked observations, or declare H_1 . Let $Y^{[m]}$ denote the LLR of rank $m = 1, 2, \dots, K$ received during the m^{th} reporting mini-slot, where $m = 1$ is the highest rank. For $1 \leq k \leq K$, the decision strategy is expressed as

$$\text{If } \sum_{m=1}^k Y^{[m]} \begin{cases} < t_k^{(L)} \\ \in (t_k^{(L)}, t_k^{(H)}) \\ > t_k^{(H)} \end{cases} \begin{matrix} \text{Declare } H_0 \\ \text{Continue} \\ \text{Declare } H_1 \end{matrix} \quad (5)$$

Note that the thresholds are generally time-dependent, i.e., their values depend on the stage index k . In order to force a stopping at $k = K$, $t_K^{(L)} = t_K^{(H)}$.

When the statistics evidently favor one hypothesis over the other, the decision is made and LLR transmissions from the local sensors stop. Otherwise, more observations are taken possibly till the end of the primary time slot. In [24], the data-dependent thresholds $t_k^{(L)}$ and $t_k^{(H)}$ are given by

$$\begin{aligned} t_k^{(L)} &= \log \left(\frac{\pi_0}{1 - \pi_0} \right) - (M - k) |Y^{[k]}| \\ t_k^{(H)} &= \log \left(\frac{\pi_0}{1 - \pi_0} \right) + (M - k) |Y^{[k]}| \end{aligned} \quad (6)$$

where $M - k$ is the number of sensors that have not transmitted their LLR's yet at the k^{th} stage, and $|Y^{[k]}|$ is the absolute value of the LLR received at stage k . However, if this scheme is forced to take a decision in favor of either H_0 or H_1 at stage K , then it should work only for the case $K = M$, as can be seen by the fact that only $t_M^{(L)} = t_M^{(H)}$. The transmission is ordered in terms of the absolute LLR value. That is, $|Y^{[m]}| < |Y^{[k]}|, \forall m > k$. Hence, $|Y^{[k]}|$ is greater than any LLR value transmitted later to the fusion center. It can be shown that this choice of thresholds ensure that this scheme has the same average probability of error as maximum a posteriori (MAP) procedure where all the LLR's are summed and compared to $\log(\pi_0/(1 - \pi_0))$ [24]. The average probability of error when the guessed hypothesis, \hat{H} , is not the true one is given by

$$\begin{aligned} P_e &= (1 - \pi_0) Pr(\hat{H} = H_0 | H_1) \\ &\quad + \pi_0 Pr(\hat{H} = H_1 | H_0) \end{aligned} \quad (7)$$

The advantage of this system is that the average number of transmissions needed to reach a decision is about half the total number of the sensors communicating with the fusion center [24].

Here, and in our ICASSP paper [31], we consider extending the scheme in [24] to work for $K \leq M$ case such that we only process the K LLR's with highest magnitudes out of the M LLR's. Given the sequence of observations $Y^{[1]} = y_1, Y^{[2]} = y_2, \dots, Y^{[K]} = y_K$, the optimal MAP block detector when the highest in magnitude K out of M LLR values are used has the following decision rule:

$$\log \left[\frac{f_{Y^{[1]..Y^{[K]}}|H_1}(y_1, y_2, \dots, y_K | H_1)}{f_{Y^{[1]..Y^{[K]}}|H_0}(y_1, y_2, \dots, y_K | H_0)} \right] \underset{H_0}{\overset{H_1}{\gtrless}} \log \frac{\pi_0}{1 - \pi_0} \quad (8)$$

where $f_{Y^{[1]..Y^{[K]}}|H}(y_1, y_2, \dots, y_K | H)$ is the joint density function of the magnitude-ordered LLR's given hypothesis H .

In this section, we consider only the case of identical sensors. Let the probability density function of LLR value y under hypothesis H be given by $f_Y(y|H)$. Assuming the statistical independence of LLR's at various sensors, the joint density functions in (8) can be given by [32]

$$\begin{aligned} f_{Y^{[1]..Y^{[K]}}|H}(y_1, y_2, \dots, y_K | H) &= \\ f_Y(y_1 | H) \dots f_Y(y_K | H) (\Pr\{|Y| \leq |y_K|\})^{M-K} \end{aligned} \quad (9)$$

(See Appendix A for details).

In order to perform sequential detection, we compute the accumulated sum at the fusion center at stage k , which is given by $\sum_{m=1}^k \log \left[\frac{f_Y(y_m | H_1)}{f_Y(y_m | H_0)} \right]$. However, this accumulated

sum can be replaced by $\sum_{m=1}^k y_m$, because as shown in Appendix A such processing will yield the same value of the LLR received at each stage k . This accumulated sum is then compared to two thresholds, $\hat{t}_k^{(L)}$ and $\hat{t}_k^{(H)}$, such that when it is greater than $\hat{t}_k^{(H)}$, H_1 is declared, and if it is smaller than $\hat{t}_k^{(L)}$, then H_0 is declared. Otherwise, if it does not cross either thresholds, the fusion center continues to accumulate the LLR's from the next ranked sensors. For all the stages with the exception of the K^{th} stage, i.e., for $1 \leq k \leq K-1$, these thresholds are given by

$$\begin{aligned}\hat{t}_k^{(L)} &= \log \frac{\pi_0}{1 - \pi_0} \\ &\quad - (K - k) |y_k| - (M - K) \max_{0 \leq y \leq |y_k|} \rho(y) \\ \hat{t}_k^{(H)} &= \log \frac{\pi_0}{1 - \pi_0} \\ &\quad + (K - k) |y_k| - (M - K) \min_{0 \leq y \leq |y_k|} \rho(y)\end{aligned}\quad (10)$$

where for LLR Y

$$\begin{aligned}\rho(y) &= \log \frac{\Pr\{|Y| \leq y | H_1\}}{\Pr\{|Y| \leq y | H_0\}} \\ &= \log \frac{\int_{-y}^y f_Y(w | H_1) dw}{\int_{-y}^y f_Y(w | H_0) dw}\end{aligned}\quad (11)$$

is a correction term to (6) that appears at each stage k to account for the probability that all the other $(M - K)$ sensors would have LLR values less than $|y_k|$. Since it is not a monotonic function, we need to span over the range from $0 < y < |y_k|$ and choose the maximum and the minimum values of $\rho(y)$. At the end of the slot, i.e., $k = K$,

$$\hat{t}_K^{(H)} = \hat{t}_K^{(L)} = \log \frac{\pi_0}{1 - \pi_0} - (M - K)\rho(y_K) \quad (12)$$

We now prove that using the thresholds in (10) ensures that the average probability of error achieved by the sequential scheme using the above thresholds is the same as the minimum average probability of error attained by a MAP detector operating on K out of M measurements with the highest absolute magnitude. Consider the situation in which the accumulated LLR exceeds $\hat{t}_k^{(H)}$.

$$\sum_{m=1}^k y_m \geq \hat{t}_k^{(H)} \quad (13)$$

Note that

$$(K - k) |y_k| \geq \sum_{m=k+1}^K |y_m| \geq - \sum_{m=k+1}^K y_m$$

Since $|y_k| \geq |y_K|$, minimizing over the domain $0 \leq y \leq |y_k|$ achieves at least the same minimum when the domain $0 \leq$

$y \leq |y_K|$ is used. Therefore,

$$\begin{aligned}\min_{0 \leq y \leq |y_k|} \rho(y) &= \min_{0 \leq y \leq |y_k|} \log \frac{\Pr\{|Y| \leq y | H_1\}}{\Pr\{|Y| \leq y | H_0\}} \\ &\leq \min_{0 \leq y \leq |y_K|} \log \frac{\Pr\{|Y| \leq y | H_1\}}{\Pr\{|Y| \leq y | H_0\}} \\ &\leq \log \frac{\Pr\{|Y| \leq |y_K| | H_1\}}{\Pr\{|Y| \leq |y_K| | H_0\}} \\ &\quad - (M - K) \min_{0 \leq y \leq |y_K|} \rho(y) \geq \\ &\quad - (M - K) \log \frac{\Pr\{|Y| \leq |y_K| | H_1\}}{\Pr\{|Y| \leq |y_K| | H_0\}}\end{aligned}$$

Hence, when $\sum_{m=1}^k y_m$ exceeds $\hat{t}_k^{(H)}$, it also exceeds

$$\log \frac{\pi_0}{1 - \pi_0} - \sum_{m=k+1}^K y_m - (M - K) \log \frac{\Pr\{|Y| \leq |y_K| | H_1\}}{\Pr\{|Y| \leq |y_K| | H_0\}}$$

This implies that

$$\sum_{m=1}^K y_m + (M - K) \log \frac{\Pr\{|Y| \leq |y_K| | H_1\}}{\Pr\{|Y| \leq |y_K| | H_0\}} \geq \log \frac{\pi_0}{1 - \pi_0}$$

which is the MAP rule for deciding H_1 based on a block of best K among M sensors. The case of deciding H_0 can be proved similarly. The benefit reaped from using the sequential scheme is that the average number of observations needed to make a decision is close to $\frac{K}{2}$, which has been showed via simulations. In the next section, we adopt a method for determining the thresholds that allows a tradeoff between the average probability of error and the speed of reaching a decision concerning primary activity.

IV. PROPOSED RANKED SEQUENTIAL SCHEME

In this section, we use dynamic programming to calculate the thresholds for sequential detection. Specifically, we employ the backward induction technique which provides the optimal action to be taken in order to minimize the overall decision cost [33]. In the analysis below we assume that the fusion center knows the statistics of all detectors, which are updated periodically over the common control channel and change relatively slowly over time as mentioned in Section II.

Since we adopt the ordered transmission scheme and assume perfect reporting, if an LLR value is received, all subsequent LLR's would have values with a lesser magnitude. In other words, if the value of the LLR received at stage $k-1$ is equal to α , then the probability $\Pr(|Y^{[m]}| > |\alpha|) = 0, \forall m \geq k$. Recall that Y_m is the LLR from the m^{th} sensor, $1 \leq m \leq M$, whereas $Y^{[m]}$ is the LLR with the m^{th} highest magnitude that is transmitted to the fusion center at the m^{th} stage with $m = 1, 2, \dots, K$.

Let π_k be the probability of the channel being idle at stage k given the sequence of observations $Y^{[1]} = y_1, Y^{[2]} = y_2, \dots, Y^{[k]} = y_k$. Probability π_k can be obtained recursively as follows:

$$\begin{aligned} \pi_k &= \Pr(H_0 | y_1, y_2, \dots, y_k) = \\ &= \frac{f_{Y^{[k]}|Y^{[1]}\dots Y^{[k-1]}, H_0}(y_k | y_1 \dots y_{k-1}, H_0) \Pr(H_0 | y_1 \dots y_{k-1})}{\sum_{r=0,1} f_{Y^{[k]}|Y^{[1]}\dots Y^{[k-1]}, H_r}(y_k | y_1 \dots y_{k-1}, H_r) \Pr(H_r | y_1 \dots y_{k-1})} \\ &= \frac{f_{Y^{[k]}|Y^{[k-1]}, H_0}(y_k | y_{k-1}, H_0) \pi_{k-1}}{\sum_{r=0,1} f_{Y^{[k]}|Y^{[k-1]}, H_r}(y_k | y_{k-1}, H_r) (r + (1-2r) \pi_{k-1})} \end{aligned} \quad (14)$$

where in the last step, we use the Markovian property induced by ordered transmissions that $Y^{[k]}$ is independent of $Y^{[1]}, Y^{[2]}, \dots, Y^{[k-2]}$ given $Y^{[k-1]}$ and either hypothesis (Theorem 2.4.3, [32]).

What is needed for backward induction is the conditional probability of $Y^{[m]}$ given $Y^{[m-1]}$ under both H_0 and H_1 . Assume that the probability density function of LLR value Y_k under hypothesis H (either H_0 or H_1) is given by $f_{Y_k}(y|H)$, which is a scaled and shifted chi-square distribution with N degrees of freedom as mentioned in Section II. Define $\beta_{k,b,H} = \Pr\{|Y_k| > |b| \mid H\}$. The value of $\beta_{k,b,H}$ can be readily computed knowing the likelihood functions of Y_k . For non-identical sensors, the joint distribution of $Y^{[m]}$ and $Y^{[m-1]}$ conditioned on hypothesis H when $-\alpha < \gamma < \alpha$ and $m \geq 2$ is given by:

$$\begin{aligned} f_{Y^{[m]}, Y^{[m-1]}}(\alpha, \gamma | H) &= \sum_{k=1}^M \sum_{j \neq k}^M f_{Y_k}(\alpha | H) f_{Y_j}(\gamma | H) \cdot \\ &\quad \sum_{S_{k,j}^{m-2}} \prod_{v \neq k,j} (\beta_{v,\gamma,H})^{x_v} (1 - \beta_{v,\alpha,H})^{(1-x_v)} \end{aligned} \quad (15)$$

where $S_{k,j}^{m-2}$ is a set of $m-2$ sensors chosen from the M sensors with sensors k and j excluded. The number of sets $S_{k,j}^{m-2}$ is $\binom{M-2}{m-2}$. Parameter x_v is equal to unity when $v \in S_{k,j}^{m-2}$ and zero otherwise. $f_{Y^{[m]}, Y^{[m-1]}}(\alpha, \gamma | H) = 0$ if $|\alpha| > |\gamma|$. Note that (15) is derived as follows. The joint distribution between $Y^{[m]}$ and $Y^{[m-1]}$ involves a summation over all possible sensor pairs. Since, given a particular hypothesis, the LLR's from the local sensors are independent, we get the product $f_{Y_k}(\alpha | H) f_{Y_j}(\gamma | H)$. For $m \geq 2$ and due to the ordered nature of LLR transmission, we have $m-2$ sensors with absolute LLR values exceeding $Y^{[m-1]} = \gamma$. The rest of the sensors should have absolute LLR values that are less than $Y^{[m]} = \alpha$. In the case of identical sensors, $f_{Y_k}(y | H) = f_Y(y | H)$ and $\beta_{k,b,H} = \beta_{b,H}$,

$$f_{Y^{[m]}, Y^{[m-1]}}(\alpha, \gamma | H) = M(M-1) \cdot$$

$$f_Y(\alpha | H) f_Y(\gamma | H) \binom{M-2}{m-2} (\beta_{\gamma,H})^{m-2} (1 - \beta_{\alpha,H})^{M-m} \quad (16)$$

In order to obtain the conditional distribution $f_{Y^{[m]}|Y^{[m-1]}}(\alpha|\gamma, H)$, we need the distribution $f_{Y^{[m-1]}}(\gamma|H)$

because

$$f_{Y^{[m]}|Y^{[m-1]}}(\alpha|\gamma, H) = \frac{f_{Y^{[m]}, Y^{[m-1]}}(\alpha, \gamma | H)}{f_{Y^{[m-1]}}(\gamma | H)} \quad (17)$$

Distribution $f_{Y^{[m-1]}}(\gamma|H)$ can be readily obtained as

$$\begin{aligned} f_{Y^{[m-1]}}(\gamma|H) &= \\ &= \sum_{k=1}^M f_{Y_k}(\gamma|H) \sum_{S_k^{m-2}} \prod_{v \neq k} (\beta_{v,\gamma,H})^{x_v} (1 - \beta_{v,\gamma,H})^{(1-x_v)} \end{aligned} \quad (18)$$

where S_k^{m-2} is a set of $m-2$ sensors chosen from the M sensors with sensor k excluded. The number of sets S_k^{m-2} is $\binom{M-1}{m-2}$. Parameter x_v is equal to unity when $v \in S_k^{m-2}$ and zero otherwise. In the case of identical sensors, $\beta_{k,\gamma,H} = \beta_{\gamma,H}$,

$$\begin{aligned} f_{Y^{[m-1]}}(\gamma|H) &= \\ M f_Y(\gamma|H) \binom{M-1}{m-2} (\beta_{\gamma,H})^{m-2} (1 - \beta_{\gamma,H})^{M-m+1} \end{aligned} \quad (19)$$

We construct now the cost-to-go function which is used to obtain the optimal policy in optimal stopping problems [34]. The optimal policy here is comprised of the optimal thresholds to be used at each stage to determine whether to stop and make a decision regarding the channel occupancy, or continue obtaining more LLR measurements. Assume that the cost of deciding i at stage k when j is the true hypothesis is λ_{ij}^k , and the cost to continue taking observations is $c \geq 0$. Define $J_k^{(K)}$ as the minimum cost-to-go at stage k of the finite horizon K .

$$\begin{aligned} J_k^{(K)}(\pi_k, Y^{[k]}) &= \min \left\{ \lambda_{00}^k \pi_k + \lambda_{01}^k (1 - \pi_k), \right. \\ &\quad \lambda_{10}^k \pi_k + \lambda_{11}^k (1 - \pi_k), \\ &\quad \left. c + \mathbb{E}_{Y^{[k+1]}|Y^{[k]}} \left[J_{k+1}^{(K)}(\pi_{k+1}, Y^{[k+1]}) | Y^{[k]} \right] \right\} \end{aligned} \quad (20)$$

where π_k and π_{k+1} are related through the expression in (14) with k replaced by $k+1$. The first term in (20) is the cost of deciding H_0 , whereas the second term is the cost of deciding H_1 . These correspond to the channel decided to be free or busy, respectively. The third term in (20) is interpreted as the expected cost when the fusion center decides that it should continue taking more observations from local sensors. The conditional expectation is given by

$$\begin{aligned} \mathbb{E}_{Y^{[k+1]}|Y^{[k]}} \left[J_{k+1}^{(K)}(\pi_{k+1}, Y^{[k+1]}) | Y^{[k]} \right] &= \\ \int_{-|Y^{[k]}|}^{|Y^{[k]}|} J_{k+1}^{(K)}(\pi_{k+1}, y) f_{Y^{[k+1]}|Y^{[k]}}(y | Y^{[k]}) dy \end{aligned} \quad (21)$$

and

$$\begin{aligned} f_{Y^{[k+1]}|Y^{[k]}}(y | Y^{[k]}) &= f_{Y^{[k+1]}|Y^{[k]}, H_0}(y | Y^{[k]}, H_0) \pi_k \\ &\quad + f_{Y^{[k+1]}|Y^{[k]}, H_1}(y | Y^{[k]}, H_1) (1 - \pi_k) \end{aligned} \quad (22)$$

The parameter c represents the tradeoff between the time taken till a decision is made and the average probability of error. As c increases, the fusion center becomes more likely to favor one of the two hypotheses in a shorter time using only a few LLR's from the cognitive detectors. At the last stage, i.e., when $k = K$, the fusion center has two choices only; to declare H_0 or H_1 . This allows backward induction from the last stage

$$J_K^{(K)}(\pi_K, Y^{[K]}) = \min \left\{ \lambda_{00}^K \pi_K + \lambda_{01}^K (1 - \pi_K), \right. \\ \left. \lambda_{10}^K \pi_K + \lambda_{11}^K (1 - \pi_K) \right\} \quad (23)$$

Note that there is no actual dependence of $J_K^{(K)}$ on $Y^{[K]}$ because there is no more sampling after the K^{th} stage. $J_{K-1}^{(K)}(\pi_{K-1}, Y^{[K-1]})$ can be obtained using (20) for all values of $Y^{[K-1]}$. The process can then be repeated to obtain the thresholds and optimal decisions. It is clear that this is a computationally extensive task due to the dependence on last observation. In particular, the thresholds at each stage would be a function of the value of the LLR sent in the previous stage. To reduce the complexity, we propose here to use the distributions of $Y^{[k]}$ without conditioning. This makes the problem considerably more manageable. We have verified through simulations that at least for a few number of sensors dropping this dependency has a negligible effect on the objective function. Dropping the explicit dependence on $Y^{[k]}$, the dynamic program now acquires the form

$$J_k^{(K)}(\pi_k) = \min \left\{ \lambda_{00}^k \pi_k + \lambda_{01}^k (1 - \pi_k), \right. \\ \left. \lambda_{10}^k \pi_k + \lambda_{11}^k (1 - \pi_k), \right. \\ \left. c + \mathbb{E}_{Y^{[k+1]}} \left[J_{k+1}^{(K)}(\pi_{k+1}) \right] \right\} \quad (24)$$

such that

$$\mathbb{E}_{Y^{[k+1]}} \left[J_{k+1}^{(K)}(\pi_{k+1}) \right] = \int J_{k+1}^{(K)}(\pi_{k+1}) f_{Y^{[k+1]}}(y) dy \quad (25)$$

with

$$f_{Y^{[k+1]}}(y) = \pi_k f_{Y^{[k+1]}}(y | H_0) + (1 - \pi_k) f_{Y^{[k+1]}}(y | H_1) \quad (26)$$

$$\pi_{k+1} = \frac{\pi_k f_{Y^{[k+1]}}(y | H_0)}{\pi_k f_{Y^{[k+1]}}(y | H_0) + (1 - \pi_k) f_{Y^{[k+1]}}(y | H_1)} \quad (27)$$

Finally, $f_{Y^{[k+1]}}(y | H)$ can be computed using an expression similar to (28)

$$f_{Y^{[k+1]}}(y | H) = \sum_{r=1}^M f_{Y_r}(y | H) \sum_{S_r^k} \prod_{v \neq r} (\beta_{v,y,H})^{x_v} (1 - \beta_{v,y,H})^{(1-x_v)} \quad (28)$$

Note that the optimal thresholds and decisions can be calculated offline so long as the relevant system parameters are fixed. When the fusion center gets the LLR observations, it uses the pre-computed thresholds to decide whether to decide in favor of H_0 or H_1 , or to request the transmission of one

more LLR value from sensors. The calculation should be redone when any of the parameters, such as the local SNR's at the cognitive detectors, change significantly.

V. PERFORMANCE OPTIMIZATION

We now provide expressions for costs λ_{ij}^k introduced in Section IV. Recall that λ_{ij}^k is the cost to decide i when j is true at the k^{th} mini-slot of duration τ . Our objective is to optimize the system performance with a focus on the achievable primary and secondary throughput. Specifically, when we deal with throughput, we mean a weighted sum of primary and secondary throughputs where a factor ω , $0 \leq \omega \leq 1$, is multiplied by primary throughput and $1 - \omega$ is used to weight the secondary throughput. The closer ω to unity, the more emphasis we put on the primary rate. This implies more protection for the primary link against interference and service interruption. Parameter ω should be chosen by the primary network to satisfy the transmission and rate requirements of primary users.

Define η_p and η_s to be the probability of correct reception of the primary and secondary signals in the presence of receiver noise only, respectively. Similarly, define δ_p and δ_s to be the probability of correct reception of the primary and secondary signals in the presence of noise in addition to interference from the other user due to concurrent transmission. Given that R_p and R_s are the primary and secondary rates of transmission, respectively, the costs of different decisions are then given by

$$\lambda_{00}^k = - (1 - \omega) R_s \eta_s \frac{(\tau_s - \tau_N - k\tau)}{\tau_s} + \\ e_{st} \frac{(\tau_s - \tau_N - k\tau)}{\tau_s} \quad (29)$$

where e_{st} is the transmission cost or expended energy for the secondary terminal if it transmits for the whole slot duration. This cost is given in rate units to be compatible with the first term in (29).

$$\lambda_{01}^k = - \omega R_p \delta_p - (1 - \omega) R_s \delta_s \frac{(\tau_s - \tau_N - k\tau)}{\tau_s} + \\ e_{pt} + e_{st} \frac{(\tau_s - \tau_N - k\tau)}{\tau_s} + P \quad (30)$$

where e_{pt} is the transmission cost or expended energy for the primary terminal expressed in units of rate, and P is a penalty term for collision with primary user.

$$\lambda_{10}^k = L_f \quad (31)$$

where L_f is the cost of losing the opportunity to access the spectrum when the channel is idle. This term can be positive to account for other negative consequences for remaining idle than the throughput being zero. The failure to transmit may increase the transmission delay for a delay-sensitive application, or cause packet loss due to queue overflow, etc.

$$\lambda_{11}^k = - \omega R_p \eta_p + e_{pt} + L_b \quad (32)$$

where L_b is the cost of losing the opportunity to access the spectrum when the channel is busy. All the costs e_{st} , e_{pt} , L_f , L_b and P are nonnegative.

The first term of expression (29) shows the gain of the secondary user when an idle channel is correctly detected. It depends on the weight assigned to the secondary user, the time remaining from the time slot τ_s after making the decision, and the probability of correct detection in the absence of primary interference. Similarly, the first term of (32) is the primary throughput in the case of correct detection by the secondary users. The first two terms in (30) represent the possibility of correct reception in spite of the occurrence of a collision due to mis-detection. The interference survival probabilities δ_p and δ_s would be typically small in value.

Recall that parameter $c \geq 0$ is the cost of taking one more observation. Given the aforementioned performance metrics, c can be understood as the cost, in units of rate, of the transmission phase in which the secondary fusion center solicits another LLR measurement and receives it from the appropriate local sensor.

A. Throughput Maximization

If we adopt the weighted sum throughput maximization objective, then setting all the costs, including c , to zero except for the throughput terms, declaring H_1 at any stage in the slot except for the last stage might cause losing the opportunity to access the spectrum by the cognitive radio device. Instead, the fusion center should solicit more observations, for a possible declaration of a free channel before the end of the slot, which consequently increases the normalized secondary throughput. Hence, rather than two thresholds as in classical sequential detection, we can just have one threshold and two decisions: to declare H_0 and assign one of the cognitive terminals to transmit over the probed channel, or to ask for one more observation if less than K LLR observations have been collected. Consequently, the objective cost-to-go at stage k can be expressed as

$$J_k^{(K)}(\pi_k) = \min \left\{ \lambda_{00}^k \pi_k + \lambda_{01}^k (1 - \pi_k), \right. \\ \left. \mathbb{E}_{Y_{k+1}} \left[J_{k+1}^{(K)}(\pi_{k+1}) \right] \right\} \quad (33)$$

where λ_{00}^k and λ_{01}^k are given by (29) and (30) with $e_{st} = e_{pt} = 0$. The above expression is valid for $k < K$. At the last stage where $k = K$, we have

$$J_K^{(K)}(\pi_K) = \min \left\{ \lambda_{00}^K \pi_K + \lambda_{01}^K (1 - \pi_K), \right. \\ \left. \lambda_{10}^K \pi_K + \lambda_{11}^K (1 - \pi_K) \right\} \quad (34)$$

where λ_{10}^K and λ_{11}^K are given by (31) and (32) with $k = K$, $e_{st} = e_{pt} = 0$ and $L_b = L_f = 0$. We actually can show that starting with the two-threshold case, if we set $c = 0$, $e_{st} = e_{pt} = 0$ and $L_b = L_f = 0$, the two-threshold case would converge to the one-threshold case as the lower threshold would always be set to 0 except at the very last stage. The proof is given in Appendix B. Note that if there are multiple channels the two-threshold case would not necessarily

converge to the one-threshold case because declaring H_1 would be equivalent to switching to another channel in order to search for a transmission opportunity.

VI. FADING CHANNELS BETWEEN SENSORS AND FUSION CENTER

In this section, and instead of the previously assumed perfect reporting channels, we consider fading channels between the local detectors and the fusion center. The statistical distribution of the channel gain between the i^{th} sensor and the fusion center is given by $f_{g_i}(x)$, possibly different for different sensors. We assume that the instantaneous values of the channel gains are known by the sensors and the fusion center. These values are fixed over a duration T_c which is multiples of the primary slot duration τ_s . Note that if the channel between a sensor and a fusion center is in deep fade, then even if the sensor has a reliable observation, it may not be able to transfer it correctly to the fusion center within the mini-slot duration τ . For the transmission from the i^{th} sensor to the fusion center to be decodable, the rate of transmission should be lower than the link capacity which, for a fixed and known channel gain g_i at the transmitter and additive white Gaussian noise at the receiver, is given by

$$C_i = W \log \left(1 + \frac{P_i g_i}{\sigma_f^2} \right) \quad (35)$$

where W is the reporting channel bandwidth, σ_f^2 is the noise variance at the receiver of the fusion center, and P_i is the transmitted power by the i^{th} sensor. We consider the case where a local sensor transmits with a rate r_i given by

$$r_i = W \log \left(1 + \frac{P_i g_i}{\Gamma_i \sigma_f^2} \right) \quad (36)$$

where $\Gamma_i > 1$ is the signal-to-noise ratio gap to capacity [35]. This factor accounts for the practical limitations that forces transmission at a rate below link capacity for correct reception. Assume that a number b of information bits needs to be transferred to the fusion center. This would include the quantized LLR values with high resolution to justify the neglect of quantization noise. The amount of time needed for this transfer conditioned that the fusion center would be able to decode the transmission correctly is given by b/r_i . We consider that this transmission should take no longer than τ_b which is obviously smaller than τ . In other words, the i^{th} sensor would be able to send its LLR value to the fusion center in time τ_b or less if

$$\frac{b}{W \log \left(1 + \frac{P_i g_i}{\Gamma_i \sigma_f^2} \right)} \leq \tau_b \quad (37)$$

This can be written as

$$g_i \geq \frac{\Gamma_i \sigma_f^2}{P_i} \left(2^{\frac{b}{W \tau_b}} - 1 \right) \quad (38)$$

Let \bar{g}_i be the right-hand-side of inequality (38). The probability then that a sensor would report its LLR measurement to the

fusion center is given by

$$\delta_i = \int_{\bar{g}_i}^{\infty} f_{g_i}(x) dx \quad (39)$$

Over a period of T_c only a fraction of the local detectors would participate in the reporting process to the fusion center. Since the channels are assumed to be known at the sensors, each sensor can decide whether it should participate in the next sensing epoch. The fusion center also, knowing the channels, can know the subset of sensors that would be involved in the next sensing events that span a time duration of T_c . Our analysis in the previous sections applies but with a number of sensors equal to those whose channels satisfy (38). The average number of sensors participating in the sensing is simply $\sum_{i=1}^M \delta_i$. The probability that the number of sensors participating in sensing is equal to \bar{M} , where $0 \leq \bar{M} \leq M$, is given by

$$\sum_{S_{\bar{M}}} \prod_{i=1}^M \delta_i^{x_i} (1 - \delta_i)^{1-x_i} \quad (40)$$

where $S_{\bar{M}}$ is a set of \bar{M} sensors from M sensors. The number of elements in set $S_{\bar{M}}$ is $\binom{M}{\bar{M}}$. Factor $x_i = 1$ if sensor i belongs to the set $S_{\bar{M}}$ and zero otherwise. For symmetric channels between the sensors and the fusion center, $\delta_i = \delta$ for all i and the probability of \bar{M} sensors participating is

$$\binom{M}{\bar{M}} \delta^{\bar{M}} (1 - \delta)^{M-\bar{M}} \quad (41)$$

The average number of sensors participating in the sensing and reporting process is, hence, equal to $M\delta$. This means that the performance of the network with M sensors would, on average, be equal to the performance of a network with non-fading reporting channels, albeit with $M\delta$ sensors.

VII. SIMULATION RESULTS

In this section, we investigate via numerical simulations the performance of our proposed schemes. Unless otherwise stated, the simulation parameters are as follows. The noise variance at each detector $\sigma^2 = 1$ and the transmission of each observation takes $\tau = 0.1$ units of time. The time taken at the beginning of a time slot with duration $\tau_s = 1$ to collect N observations is $\tau_N = 2\tau$. This leaves $K = 8$ stages in the slot in which the fusion center takes the ordered LLR observations. We consider the case of equal prior probabilities of both hypotheses, i.e., $\pi_0 = 0.5$, and equal local signal to noise ratio over the channels between the local sensors and the primary user, $\sigma_{s_i}^2 = 2$, $i = 1, 2, \dots, M$. Each sensor takes $N = 3$ samples in the duration τ_N . The cost, c , to continue without deciding on one of the two hypotheses is set to 0.0001.

In order to minimize the average probability of error in equation (24), we use the decision costs: $\lambda_{ij}^k = 1$ when $i \neq j$ and $\lambda_{ij}^k = 0$ when $i = j$. Cost $\lambda_{ii}^k = 0$ means that no cost is incurred if the decision at stage k is the true hypothesis. Moreover, to study the objective of maximizing the weighted sum throughput, we use the expressions (29)-(32) for the costs. In the simulations, the transmission costs for both the primary

and secondary users, e_{pt} and e_{st} , are set to 0, as well as the penalty factor P . Furthermore, the costs of losing the opportunity to access the spectrum when it is idle or busy, L_f and L_b , are assumed to be 0. The transmission success probabilities are $\eta_p = 1$, $\eta_s = 1$, $\delta_p = 0$ and $\delta_s = 0$.

A. Two-Threshold Based Method

Figure 2 shows the normalized primary and secondary throughputs for a weight $\omega = 0.5$ and a weight $\omega = 0.999$. The latter case is the case of interest in the context of cognitive radios given the privileges of the licensed primary users. The normalized throughput is obtained via simulating the system using the optimal thresholds. If the simulation is run over Q time slots, the normalized secondary throughput is given by

$$\frac{1}{Q} \sum_{q=1}^Q I_q^{(s)} R_s \left(1 - \frac{\tau_N + k_q \tau}{\tau_s} \right) \quad (42)$$

where $I_q^{(s)}$ is equal to unity if the one of the secondary users transmits successfully over the q^{th} time slot and zero otherwise, and k_q is the number of probed sensors in the q^{th} time slot. The normalized primary throughput is equal to $\frac{1}{Q} \sum_{q=1}^Q I_q^{(p)} R_p$, where $I_q^{(p)}$ is equal to unity when the primary user transmits successfully during the q^{th} time slot and zero otherwise. It is clear from the figure that increasing the weight puts more emphasis on the primary throughput. The difference is, however, small in the case of a large number of sensors. For the rest of the results in this section and just for demonstration purposes, we use $\omega = 0.5$.

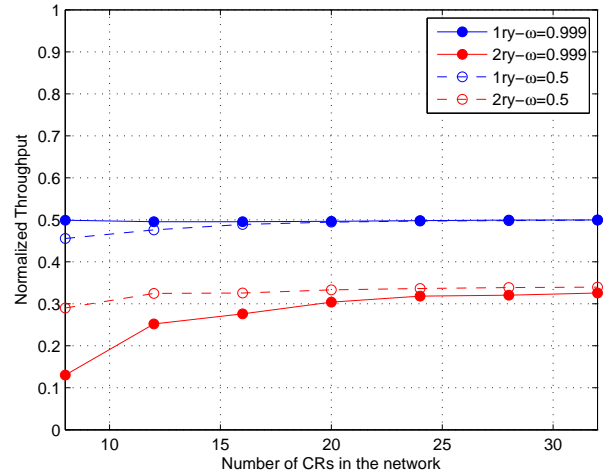


Fig. 2. Normalized primary and secondary throughputs vs. M , with $N = 3$, $\sigma_{s_i}^2 = 2$, $K = 8$, $\eta_p = 1 = \eta_s = 1$, $R_p = R_s = 1$, $\delta_p = \delta_s = 0$ and $c = 0.0001$ for $\omega = 0.5$ and $\omega = 0.999$. In the legend, “1ry” refers to primary and “2ry” refers to secondary. The scheme used to generate this figure is dynamic programming for throughput optimization.

In Figure 3, it is clear that the modified scheme of [24], which we hereafter refer to as Modified B&S scheme, achieves a lower probability of error than the dynamic programming scheme, denoted as DP scheme, since it achieves the same

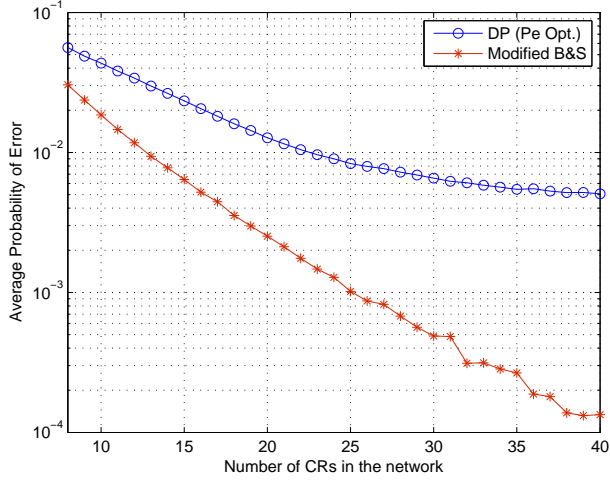


Fig. 3. Average probability of error vs. M , $N = 3$, $\sigma_{s_i}^2 = 2$, $K = 8$, and $c = 0.0001$.

probability of error as the MAP block detector. However, as illustrated in Figure 4, the number of sensors involved in the sensing process in the dynamic programming scheme is lower than that in the Modified B&S scheme. The number of sensors actually converges slowly to unity using dynamic programming. Increasing the number of sensors in the network causes the ranked LLR observations to be more informative about the status of the channel. Therefore, when the DP scheme is used, the decision can be made after taking observations from one sensor on average, given that it has the maximum absolute LLR. However, for the Modified B&S scheme, the sensing time increases as the number of CR's in the network increases. This follows from the fact that the greater the number of sensors in the network, the greater the distance between the two thresholds of comparison in (10) causing more observations to be taken. Thus, the dynamic programming scheme represents a tradeoff of the sensing time with the error probability, via the c parameter. The performance enhancement increases when the objective becomes to maximize the achievable weighted sum throughput as is evident from Figure 4. Figure 5 demonstrates these results on the normalized secondary throughput, assuming unity primary and secondary transmission rates R_p and R_s . For the normalized secondary throughput it is obvious from the simulation results that it converges to approximately $\pi_0 \left(1 - \frac{\tau_N + \tau}{\tau_s}\right)$ in the DP scheme. This is the throughput for the hypothetical case of using just one "perfect" sensor that reveals that true state of the channel.

Using Modified B&S scheme, it is noted that when the local sensors' SNR is high, the number of probed sensors is almost upperbounded by $\frac{K}{2}$ as shown in Figure 6. For low SNR regime, the number of probed sensors is close to $\frac{K}{2}$, though may exceed it when the SNR is low depending on the value of K . If we use conditional distributions for the measurements at local sensors for which the correction term $\rho(y)$ in equation (10) is zero, it can be shown analytically following a proof

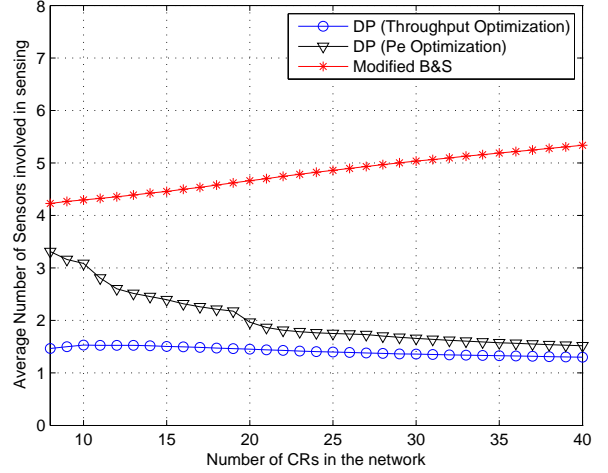


Fig. 4. Average number of probed sensors for both DP schemes and the Modified B&S scheme, vs. M , $N = 3$, $\sigma_{s_i}^2 = 2$, $K = 8$, and $c = 0.0001$.

in [24] that the maximum number of probed sensors is $\frac{K}{2}$. This is the case for example when the local measurements have conditional distributions that, in contrast with those in (2), are Gaussian with different means and the same variance. The simulation results for such "shift in mean" case are also provided in Figure 6.

Taking into consideration the fading effect in the channels between the sensors and the fusion center as explained in Section VI, only a subset of the total number of local detectors becomes involved in the sensing and reporting process according to the fading coefficients. This means that the fusion center takes longer time to detect the presence or absence of the primary user. Figure 7 shows the effect of fading, which is to slightly increase the average number of probed sensors. The parameters used to produce this figure are $b = 20$, $W = 50$ kHz, $\tau_b = \frac{\tau}{2} = 0.5$ msec, $\Gamma_i = 2$, and $\frac{P_i}{\sigma_i^2} = 5$. Assuming that the channel gains between the sensors and the fusion center are i.i.d and are exponentially distributed with unity mean, we can easily find that δ is equal to 0.743. Therefore, slightly less than three quarters of the number of sensors participate, on average, in the process of spectrum sensing. The observation that fading has a minor effect is just because of the simulation parameters. Note that in Figure 4 the average number of sensors is already less than 2 even for a relatively small M .

B. One-Threshold Based Method

As mentioned in Section V, instead of applying the two-threshold based scheme to the accumulated LLR observations at the fusion center, it can be sufficient to perform the comparison using one threshold only, when the sensing cost, c , is set to 0. In fact, this requires consuming more mini-slots in sensing the channel, till a decision is made about the channel being idle. Otherwise, sensing continues till the end of the slot, declaring a busy channel.

Figure 8 shows the convergence of the two-threshold scheme to the one-threshold scheme at zero sensing cost, for

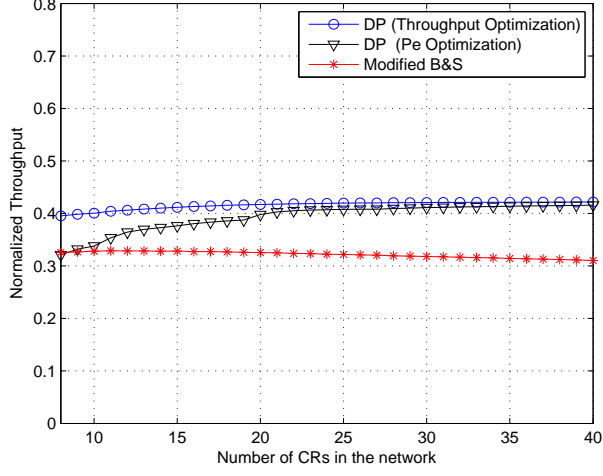


Fig. 5. Normalized Weighted Sum Throughputs for both DP schemes and the Modified B&S scheme, vs. M , $N = 3$, $\sigma_{s_i}^2 = 2$, $K = 8$, and $c = 0.0001$.

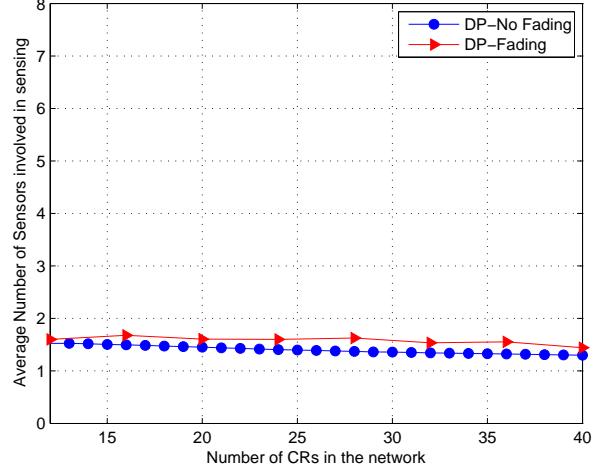


Fig. 7. Average sensing time vs. M , with $N = 3$, $\sigma_{s_i}^2 = 2$, $K = 8$, using the DP scheme for error optimization under fading between the local sensors and the fusion center.

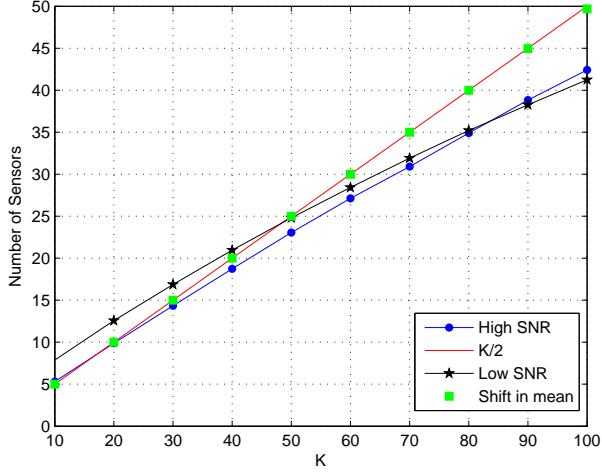


Fig. 6. Number of sensors involved in sensing, for high SNR and low SNR versus different values for K , with $M = 100$, $N = 3$ and $\sigma_{s_i}^2 = 2$ for low SNR and $\sigma_{s_i}^2 = 50$ for high SNR. The “shift in mean” case refers to using conditional distributions for local sensor measurements that, in contrast with (2), are Gaussian with the same variance and different means.

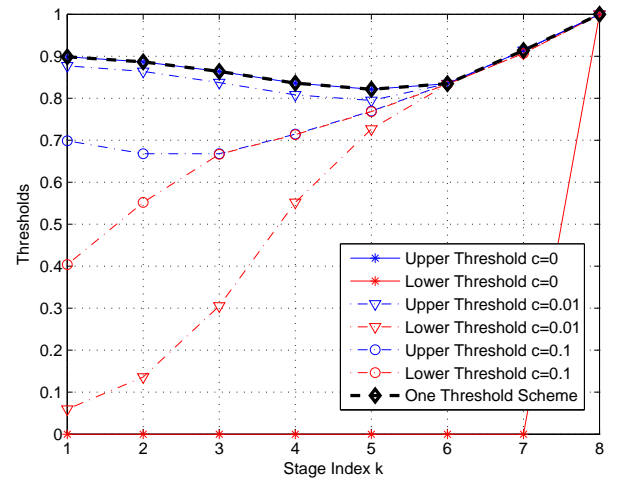


Fig. 8. Thresholds of comparison for different costs vs. stage index k , with $K = 8$, $M = 10$, $\sigma_{s_i}^2 = 2$ and $N = 3$.

the $M = 10$ case. As c decreases, the lower threshold is reduced at all stages. When $c = 0$, the lower threshold is 0 at all stages except for the last one at which a decision must be made. The fusion center therefore continues to take observations till the end of the slot if there is no sufficient statistics to declare H_0 . At any stage up to $k = 8$, the fusion center has only two choices; to declare H_0 or to probe more sensors, which is the main idea of the one-threshold scheme.

To investigate the effect of the cost parameter c on the sensing time, we consider the dynamic programming scheme with decision costs set for minimizing the average probability of error. Figure 9 shows that at $c = 0$ the fusion center does not make a decision till the end of the slot. Increasing the cost

gives the fusion center the chance to make a decision before the end of the slot allowing for secondary transmission.

VIII. CONCLUSION

In this paper, we considered ordering the transmissions from cognitive users to the decision fusion center according to the information they carry about the status of the primary user. Restricting the transmissions to the highly informative set of observations, i.e., the LLR's with the highest magnitude, increases the time remaining in the time slot after making the decision, thereby boosting the throughput while protecting the primary signal. We devised a sequential scheme that achieves the optimum error probability of the MAP block detector, but makes the decision faster. We also devised a scheme based on dynamic programming that allows a trades-off between

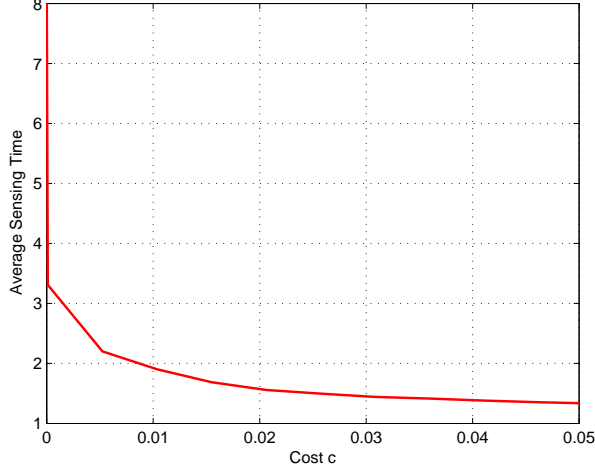


Fig. 9. Average sensing time vs. c , with $N = 3$, $\sigma_{s_i}^2 = 2$, $K = M = 8$, using the DP scheme for error optimization.

the average probability of error and the average sensing time. Simulation results have demonstrated that using dynamic programming approach to compute the thresholds required to make the decision at the fusion center yields a considerable enhancement in the performance in terms of reduced delay and increased secondary throughput.

Future work needs to address the situation of multiple primary channels. Also of interest is the design of the multiple access scheme that allows the sensor with the highest LLR in magnitude to seize the common control channel and transmit. Another front that needs further investigation is the channel access scheme and how the discovered spectrum opportunities are to be distributed among the secondary users.

APPENDIX A

For the optimal probability of error MAP detector, using the best K out of M LLR magnitudes, the joint distribution of K observations under hypothesis H (either H_0 or H_1) is given by $f_{Y_1..Y_K|H}(y_1, y_2, ..y_K | H)$. Since the sequence of observations forms a Markov chain (Theorem 2.4.3, [32])

$$f_{Y_1, Y_2, \dots, Y_K | H}(y_1, y_2, ..y_K | H) = f_{Y^{[1]}}(Y^{[1]} = y_1) \cdot \prod_{j=2}^K f_{Y^{[j]} | Y^{[j-1]}, H}(Y^{[j]} = y_j | Y^{[j-1]} = y_{j-1}, H) \quad (43)$$

For symmetric sensors, the conditional distribution of the sensor of rank j receiving an LLR value $Y^{[j]} = \alpha$ given that the sensor of rank $j - 1$ has already received an LLR value

$Y^{[j-1]} = \gamma$ is given by

$$f_{Y^{[j]} | Y^{[j-1]}, H}(Y^{[j]} = \alpha | Y^{[j-1]} = \gamma, H) = \frac{(\beta_{\gamma, H})^{j-2}}{(\beta_{\alpha, H})^{j-2}} \cdot \frac{(1 - \beta_{\alpha, H})^{M-j} M(M-1) \cdot f_Y(\alpha | H) f_Y(\gamma | H) \binom{M-2}{j-2}}{(1 - \beta_{\alpha, H})^{M-j+1} M \cdot f_Y(\gamma | H) \binom{M-1}{j-2}} = (M+1-j) f_Y(\alpha | H) \cdot \frac{(1 - \beta_{\alpha, H})^{M-j}}{(1 - \beta_{\gamma, H})^{M-j+1}} \quad (44)$$

with $\beta_{x, H}$ defined as $\Pr\{|Y| \geq |x| | H\}$, where random variable Y is an LLR observation. Hence,

$$\begin{aligned} f_{Y_1, Y_2, \dots, Y_K | H}(y_1, y_2, ..y_K | H) &= M f_Y(y_1 | H) (1 - \beta_{y_1, H})^{M-1} \cdot \\ &\quad (M-1) f_Y(y_2 | H) \frac{(1 - \beta_{y_2, H})^{M-2}}{(1 - \beta_{y_1, H})^{M-1}} \cdot \\ &\quad (M-2) f_Y(y_3 | H) \frac{(1 - \beta_{y_3, H})^{M-3}}{(1 - \beta_{y_2, H})^{M-2}} \cdot \\ &\quad \dots (M+1-K) f_Y(y_K | H) \frac{(1 - \beta_{y_K, H})^{M-K}}{(1 - \beta_{y_{K-1}, H})^{M-K+1}} \end{aligned} \quad (45)$$

Then, the likelihood ratio of the best K out of M LLR's for a block detector can be written as

$$\frac{f_{Y_1..Y_K | H_1}(y_1, y_2, ..y_K | H_1)}{f_{Y_1..Y_K | H_0}(y_1, y_2, ..y_K | H_0)} = \frac{f_Y(y_1 | H_1) \dots f_Y(y_K | H_1) (1 - \beta_{y_K, H_1})^{M-K}}{f_Y(y_1 | H_0) \dots f_Y(y_K | H_0) (1 - \beta_{y_K, H_0})^{M-K}} \quad (46)$$

We now show that the sequential algorithm may operate on the received LLR's directly rather than computing their LLR's. This requires that the LLR is a strictly monotonic function of the observation. Consider observation y with llr given by

$$\text{llr} = \log \frac{f_Y(y | H_1)}{f_Y(y | H_0)} = g(y) \quad (47)$$

For a monotonic g function, conditioned on H (either H_1 or H_0),

$$f_{LLR}(\text{llr} | H) = f_Y(y | H) \Big|_{y=g^{-1}(\text{llr})} \cdot \left| \frac{dg^{-1}(y)}{d\text{llr}} \right| \quad (48)$$

Then,

$$\log \frac{f_{LLR}(\text{llr} | H_1)}{f_{LLR}(\text{llr} | H_0)} = \log \frac{f_Y(y | H_1) \Big|_{y=g^{-1}(\text{llr})}}{f_Y(y | H_0) \Big|_{y=g^{-1}(\text{llr})}} \quad (49)$$

But the right-hand-side of (49) is given by

$$\begin{aligned} \log \frac{f_Y(y | H_1) \Big|_{y=g^{-1}(\text{llr})}}{f_Y(y | H_0) \Big|_{y=g^{-1}(\text{llr})}} &= \log \frac{f_Y(g^{-1}(\text{llr}) | H_1)}{f_Y(g^{-1}(\text{llr}) | H_0)} \\ &= g(g^{-1}(\text{llr})) \\ &= \text{llr} \end{aligned} \quad (50)$$

Therefore,

$$\log \frac{f_{LLR}(\text{llr} | H_1)}{f_{LLR}(\text{llr} | H_0)} = \text{llr} \quad (51)$$

If the LLR monotonicity assumption is not satisfied, the thresholds in (10) can be modified as follows:

$$\begin{aligned} \hat{t}_k^{(H)} &= \log \frac{\pi_0}{1 - \pi_0} - \min_{0 \leq y \leq |y_k|} \bar{\rho}(y) \\ \hat{t}_k^{(L)} &= \log \frac{\pi_0}{1 - \pi_0} - \max_{0 \leq y \leq |y_k|} \bar{\rho}(y) \end{aligned} \quad (52)$$

where

$$\bar{\rho}(y) = (M - K) \rho(y) + (K - k) \log \frac{f_Y(y | H_1)}{f_Y(y | H_0)} \quad (53)$$

The proof that these thresholds achieve the same average probability of error if the best K out of M sensors are used follows the same outline as that presented in Section (III). We focus here on the lower threshold. A decision in favor of H_0 is made if the metric $\sum_{m=1}^k \log \frac{f_Y(y_m | H_1)}{f_Y(y_m | H_0)}$ gets below $\hat{t}_k^{(L)}$. If $\sum_{m=1}^k \log \frac{f_Y(y_m | H_1)}{f_Y(y_m | H_0)} < \hat{t}_k^{(L)}$,

$$\begin{aligned} \sum_{m=1}^k \log \frac{f_Y(y_m | H_1)}{f_Y(y_m | H_0)} &< \log \frac{\pi_0}{1 - \pi_0} - \max_{0 \leq y \leq |y_k|} \bar{\rho}(y) < \\ \log \frac{\pi_0}{1 - \pi_0} - \sum_{m=k+1}^K \log \frac{f_Y(y_m | H_1)}{f_Y(y_m | H_0)} - (M - K) \max_{0 \leq y \leq |y_k|} \rho(y) \end{aligned}$$

Then

$$\begin{aligned} \sum_{m=1}^K \log \frac{f_Y(y_m | H_1)}{f_Y(y_m | H_0)} &< \log \frac{\pi_0}{1 - \pi_0} - (M - K) \max_{0 \leq y \leq |y_k|} \rho(y) \\ &< \log \frac{\pi_0}{1 - \pi_0} - (M - K) \max_{0 \leq y \leq |y_k|} \rho(y) \\ &< \log \frac{\pi_0}{1 - \pi_0} - (M - K) \log \frac{1 - \beta_{y_K, H_1}}{1 - \beta_{y_K, H_0}} \end{aligned} \quad (54)$$

which is the MAP rule for a decision based on a block of the K most reliable measurements among M observations.

APPENDIX B

In this appendix we prove that when the throughput terms are preserved while all other costs, including parameter c , are set to zero, the lower detection threshold for the sequential procedure becomes always zero except for the last stage. This means that, from a throughput point of view, the optimal decisions based on the observations are either to declare H_0 and then select a cognitive terminal for channel access, or to continue taking more samples so long as less than K LLR values have been gathered.

We start by proving the concavity of function $J_K^K(\pi_K)$ using induction. At the last stage

$$\begin{aligned} J_K^K(\pi_K) &= \min \left\{ -r_s \left(1 - \frac{\tau_N + K\tau}{\tau_s} \right) \pi_K, \right. \\ &\quad \left. -r_p (1 - \pi_K) \right\} \end{aligned} \quad (55)$$

where $r_s = (1 - \omega) R_s$ and $r_p = \omega R_p$. For $k < K$,

$$\begin{aligned} J_k^{(K)}(\pi_k) &= \min \left\{ -r_s \left(1 - \frac{\tau_N + k\tau}{\tau_s} \right) \pi_k, \right. \\ &\quad \left. -r_p (1 - \pi_k), c + \Psi_k(\pi_k) \right\} \end{aligned} \quad (56)$$

where $\Psi_k(\pi_k)$ is given by

$$\Psi_k(\pi_k) = \int J_{k+1}^K \left(\frac{\pi_k f_0}{\pi_k f_0 + (1 - \pi_k) f_1} \right) \cdot \left[\pi_k f_0 + (1 - \pi_k) f_1 \right] dY^{[k+1]} \quad (57)$$

Note that we use f_0 in place of $f_{Y^{[k+1]}}(y | H_0)$ and f_1 in place of $f_{Y^{[k+1]}}(y | H_1)$ to simplify notation. Parameter c is kept in expression (56) and is set to zero after the concavity of $J_k^K(\pi_k)$ is established.

If $\Psi_k(\pi_k)$ is concave, then $c + \Psi_k(\pi_k)$ is concave and, consequently, $J_k^K(\pi_k)$ is concave because $J_k^K(\pi_k)$ is the minimum of three concave terms [36]. We now prove the concavity of $\Psi_k(\pi_k)$ under the assumption that $J_{k+1}^K(\pi_{k+1})$ is concave. Function $\Psi_k(\pi_k)$ is concave if it satisfies the following inequality

$$\kappa \Psi_k(x) + (1 - \kappa) \Psi_k(z) \leq \Psi_k(\kappa x + (1 - \kappa) z) \quad (58)$$

$\forall \kappa \in [0, 1]$ [36]. Let

$$\epsilon = \kappa [x f_0 + (1 - x) f_1] + (1 - \kappa) [z f_0 + (1 - z) f_1] \quad (59)$$

The right-hand-side of (58) then can be written as

$$\begin{aligned} &\int \left\{ \frac{\kappa}{\epsilon} [x f_0 + (1 - x) f_1] J_{k+1}^K \left(\frac{x f_0}{x f_0 + (1 - x) f_1} \right) + \right. \\ &\quad \left. \frac{1 - \kappa}{\epsilon} [z f_0 + (1 - z) f_1] J_{k+1}^K \left(\frac{z f_0}{z f_0 + (1 - z) f_1} \right) \right\} \epsilon dY^{[k+1]} \end{aligned} \quad (60)$$

Noting that $\frac{\kappa}{\epsilon} [x f_0 + (1 - x) f_1] + \frac{1 - \kappa}{\epsilon} [z f_0 + (1 - z) f_1] = 1$ and using the assumption that J_{k+1}^K is concave, we obtain

$$\begin{aligned} \kappa \Psi_k(x) + (1 - \kappa) \Psi_k(z) &\leq \\ &\int J_{k+1}^K \left(\frac{\kappa}{\epsilon} x f_0 + \frac{1 - \kappa}{\epsilon} z f_0 \right) \epsilon dY^{[k+1]} \end{aligned} \quad (61)$$

Factor ϵ can be expressed as $\epsilon = \bar{\epsilon} f_0 + (1 - \bar{\epsilon}) f_1$, where $\bar{\epsilon} = \kappa x + (1 - \kappa) z$. Using this, inequality (61) becomes

$$\begin{aligned} \kappa \Psi_k(x) + (1 - \kappa) \Psi_k(z) &\leq \\ &\int J_{k+1}^K \left(\frac{\bar{\epsilon} f_0}{\bar{\epsilon} f_0 + (1 - \bar{\epsilon}) f_1} \right) [\bar{\epsilon} f_0 + (1 - \bar{\epsilon}) f_1] dY^{[k+1]} \\ &= \Psi_k(\bar{\epsilon}) = \Psi_k(\kappa x + (1 - \kappa) z) \end{aligned} \quad (62)$$

Therefore, $\Psi_k(\pi_k)$ is concave in π_k assuming that $J_{k+1}^K(\pi_{k+1})$ is concave. It is evident from (55) that $J_K^K(\pi_K)$ is concave since it is the minimum of two affine terms. By induction both $\Psi_k(\pi_k)$ and $J_k^K(\pi_k)$ are concave in π_k for $k < K$.

We now study the value of $J_k^K(\pi_k)$ at $\pi_k = 0$. It is clear that $J_K^K(0) = -r_p$. Since $\Psi_{K-1}(0) = \int J_K^K(0) f_1 dY^{[K]} = -r_p$,

then $J_{K-1}^K(0) = \min\{0, -r_p, c - r_p\} = -r_p$ for $c \geq 0$. It is straightforward to show that $\Psi_k(0) = J_k^K(0) = -r_p$ for $1 \leq k < K$.

When $\pi_k = 0$, the minimum of the three terms of (56) is $-r_p$ obtained from the second term $-r_p(1 - \pi_k)$ for $c > 0$. The value of the derivative of $J_k^K(\pi_k)$ at $\pi_k = 0$ is equal to the derivative of the term $-r_p(1 - \pi_k)$ at $\pi_k = 0$, which is equal to r_p . The derivative of $\Psi_k(\pi_k)$ is given by

$$\frac{d\Psi_k}{d\pi_k} = \int \frac{dJ_{k+1}^K}{d\pi_{k+1}} \frac{d\pi_{k+1}}{d\pi_k} [\pi_k f_0 + (1 - \pi_k) f_1] dY^{[k+1]} + J_{k+1}^K(\pi_{k+1}) [f_0 - f_1] dY^{[k+1]} \quad (63)$$

since $\pi_{k+1} = \frac{\pi_k f_0}{\pi_k f_0 + (1 - \pi_k) f_1}$, $\pi_{k+1} = 0$ when $\pi_k = 0$ and it is straightforward to show the $\frac{d\pi_{k+1}}{d\pi_k}$ at $\pi_k = 0$ is equal to $\frac{f_0}{f_1}$. Hence,

$$\left. \frac{d\Psi_k}{d\pi_k} \right|_{\pi_k=0} = \int \left(r_p \frac{f_0}{f_1} f_1 - r_p [f_0 - f_1] \right) dY^{[k+1]} = r_p \quad (64)$$

Since Ψ_k is concave, it satisfies the following inequality [36]

$$\Psi_k(z) \leq \Psi_k(x) + \left. \frac{d\Psi_k}{dx} \right|_x (z - x) \quad (65)$$

For $x = 0$ and $z = \pi_k$, $\Psi_k(\pi_k) \leq -r_p + r_p\pi_k = -r_p(1 - \pi_k)$. This means that when $c = 0$, the third term in (56) is always less than or equal the second term $-r_p(1 - \pi_k)$. Hence, the cost of getting one more sample is always less than or equal to the cost of declaring H_1 . In other words, when $c = 0$, the lower decision threshold is zero at all the stages with the exception of the last one because it is less costly to continue sampling rather than declaring H_1 .

REFERENCES

- [1] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE J. Select. Areas in Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [2] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, "Collaborative wideband sensing for cognitive radios," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 60–73, November 2008.
- [3] Ian F. Akyildiz, Brandon F. Lo, and Ravikumar Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," *Physical Communication*, vol. 4, no. 1, pp. 40–62, 2011.
- [4] G. Ganesan and Y. Li, "Cooperative Spectrum Sensing in Cognitive Radio Networks," in *Proc. IEEE DySPAN*, Nov. 2005, pp. 137–143.
- [5] A. Ghasemi and E. Sousa, "Collaborative Spectrum Sensing for Opportunistic Access in Fading Environments," in *Proc. IEEE DySPAN*, Nov. 2005, pp. 131–136.
- [6] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative Sensing Among Cognitive Radio," in *Proc. IEEE International Conference on Communications (ICC)*, Istanbul, Turkey, June 11–14, 2006, pp. 1658–1663.
- [7] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE J. Select. Topics in Signal Processing*, vol. 2, no. 1, pp. 28–40, February 2008.
- [8] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Signal Processing*, Vol. 2, H. V. Poor and J. B. Thomas, Eds., vol. 2, pp. 297–344. JAI Press, 1993.
- [9] A. Wald, *Sequential Analysis*, John Wiley and Sons, New York, 1947.
- [10] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 2nd edition, 1994.
- [11] C. W. Helstrom, *Elements of Signal Detection and Estimation*, Prentice Hall, Englewood Cliffs, N.J., 1995.
- [12] D. Teneketzis and Y. C. Ho, "The Decentralized Wald Problem," *Information and Computation*, vol. 73, no. 1, pp. 23–44, April 1987.
- [13] M. M. Al-Ibrahim and P. K. Varshney, "A simple multi-sensor sequential detection procedure," in *Proc. IEEE 27th Conference on Decision and Control*, Austin, Texas, December 1988, pp. 2479–2483.
- [14] J. N. Tsitsiklis, "On threshold rules in decentralized detection," in *Proc. IEEE 25th Conference on Decision and Control*, 1986, pp. 232–236.
- [15] H. R. Hashemi and I. B. Rhodes, "Decentralized sequential detection," *IEEE Trans. Info. Theory*, vol. 35, no. 3, pp. 509–520, May 1989.
- [16] V. V. Veeravalli, T. Basar, and H. V. Poor, "Decentralized sequential detection with a fusion center performing the sequential test," *IEEE Trans. Info. Theory*, vol. 39, no. 2, pp. 433–442, March 1993.
- [17] V. V. Veeravalli, T. Basar, and H. V. Poor, "Decentralized sequential detection with sensors performing sequential tests," *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 7, no. 4, pp. 292–305, December 1994.
- [18] R. S. Blum, S. A. Kassam, and H. V. Poor, "Distributed Detection with Multiple Sensors: Part II - Advanced Topics," *Proc. IEEE*, vol. 85, no. 1, pp. 6479, Jan. 1997.
- [19] Q. Cheng, P.K. Varshney, K.G. Mehrotra, and C.K. Mohan, "Bandwidth management in distributed sequential detection," *Information Theory, IEEE Transactions on*, vol. 51, no. 8, pp. 2954–2961, Aug. 2005.
- [20] N. Kundargi and A. Tewfik, "Hierarchical sequential detection in the context of dynamic spectrum access for cognitive radios," in *Proc. IEEE 14th Conference on Electronics, Circuits and Systems*, 2007.
- [21] Q. Zou, S. Zheng, and A. H. Sayed, "Cooperative spectrum sensing via sequential detection for cognitive radio networks," in *IEEE 10th Workshop on Signal Processing Advances in Wireless Communications (SPAWC'09)*, Perugia, Italy, June 21–24, 2009, pp. 121–125.
- [22] Y. Rao, W. Chen, and Z. Cao, "A sequential sensing data transmission and fusion approach for large scale cognitive radios," in *Proc. IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, May 23–27, 2010.
- [23] N. Kundargi and A. Tewfik, "Doubly sequential energy detection for distributed dynamic spectrum access," in *Proc. IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, May 23–27, 2010.
- [24] R. S. Blum and B.M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmission," *IEEE Trans. Signal Proc.*, vol. 56, no. 7, pp. 3229–3235, July 2008.
- [25] Yingwei Yao, "Group-ordered sprt for distributed detection," in *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'08*, March 2008, pp. 2525–2528.
- [26] Y-R. Tsai and Li-C. Lin, "Sequential fusion for distributed detection over bsc channels in an inhomogeneous sensing environment," *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 99102, Jan. 2010.
- [27] S.-J. Kim and G. B. Giannakis, "Sequential and cooperative sensing for multi-channel cognitive radios," *IEEE Trans. Signal Proc.*, vol. 58, no. 8, pp. 4239–4253, August 2010.
- [28] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, April 2008.
- [29] Nhan Nguyen-Thanh and Insoo Koo, "Log-likelihood ratio optimal quantizer for cooperative spectrum sensing in cognitive radio," *IEEE Communications Letters*, vol. 15, no. 3, pp. 317–319, March 2011.
- [30] R. S. Blum, "Ordering for estimation and optimization in energy efficient sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2847–2856, June 2011.
- [31] L. Hesham, A. Sultan, M. Nafie, and F. Digham, "Cooperative Sensing with Sequential Ordered Transmissions to Secondary Fusion Center," *IEEE ICASSP*, May 2011.
- [32] Barry C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *A First Course in Order Statistics*, John Wiley and Sons, Philadelphia, PA, USA, July 1992.
- [33] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1 and 2, Athena Scientific, 2nd edition, 2001.
- [34] T. S. Ferguson, *Optimal Stopping and Applications*, online book. Available at <http://www.math.ucla.edu/tom/Stopping/Contents.html>. Last accessed on May 24, 2011.
- [35] Edward A. Lee and David G. Messerschmitt, *Digital Communication: Second Edition*, chapter 8, Kluwer Academic Publishers, 1994.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.